

Amendment  
Serial No. 10/076,194

Docket No. US020037

### REMARKS

Entry of this Amendment and reconsideration are respectfully requested in view of the amendments made to the claims and for the remarks made herein.

Claims 1, 2, and 4-20 are pending and stand rejected.

Claims 1, 2, 4, 5, 8, 11 and 16-17 stand rejected under 35 USC 103(a) as being unpatentable over Basu (USP no. 6,219,640) in view of Nevenka (USPPA 2003/0108334).

Applicant respectfully disagrees with and explicitly traverses the reason for rejecting the claims.

Basu describes in one aspect a method of performing a face recognition step and audio recognition step and combining the individually determined face and audio values to determine a match of video image to audio. See, for example, Figure 1, and col. 8, lines 25-52, which state in part, "[n]ext, the results of the face recognition module 24 and the audio speaker recognition module 16 are provided to respective confidence estimation blocks 26 and 18 where confidence estimation is performed... Given the audio-based speaker recognition and face recognition scores provided by respective modules 16 and 14, audio-visual speaker identification/verification may be performed by a joint identification/verification module 30 as follows. The top N scores are generated-based on both audio and video-based identification techniques. The two lists are combined by a weighted sum and the best-scoring candidate is chosen..."

Basu further discloses an "alternative embodiment of an audio-visual speaker recognition and utterance verification system is shown [in Figure 3]. Whereas the embodiment of FIG. 1 illustrates a decision or score fusion approach, the embodiment of FIG. 3 illustrates a feature fusion approach. The operations of the system of FIG. 3 are substantially the same as those described above with respect to FIG. 1, however, the embodiment of FIG. 3 has the added advantage [of] making an identification/verification decision on a combined AV feature vector. In accordance with the feature fusion approach, a single feature vector is built combining acoustic features (e.g., mel cepstra and derivatives) from the acoustic feature extractor 14 and detected visual facial features ... The features are combined to form a single audio-visual feature vector. ...[T]here is a

Amendment  
Serial No. 10/076,194

Docket No. US020037

need for synchronization between feature. ... Examples are linear interpolation from frames immediately preceding and following the time instant or other polynomial interpolation techniques ... A decision operation such as, for example, that described above with respect to module 30 in FIG. 1 is performed on the combined audio-visual feature vector."

Accordingly, in the alternative embodiment, Basu describes a system that forms a single AV vector for a frame at a time instant and frames each of the time instant before and after the time instant and then uses a correlation to determine that AV vector that best describes the facial/audio features at the time instant.

Hence, in the embodiment shown in Figure 1, Basu discloses separate facial and audio determinations are performing a correlation among the N best scores of each to choose that combination with the highest score. In the embodiment of Figure 3, Basu discloses forming a single AV vector for each of the period around a time instant and performing a correlation to choose the vector with the highest score.

Basu fails to suggest or describe determining "a maximum correlation value among a plurality of correlation values between the plurality of object features and the plurality of audio features, wherein said correlation values are determined as the sum elements in a subset of said audio features selected from the group consisting of: two or more of the following: average energy, pitch, zero crossing, bandwidth, band central, roll off, low ratio, spectral flux and 12 MFCC components, and selected object features," as is recited in the claims. Rather, as Basu describes the correlation is performed, in Figure 3, for each of the single AV vectors created for each frame. Nowhere does Basu describe determining the AV vector as that vector having a maximum correlation value wherein the correlation values are determined from a subset of the audio and video features.

Nevenka is recited to show that audio elements may be composed of low level elements of bandwidth, energy and pitch. However, Nevenda fails to describe creating AV vectors from a subset of the audio and video features and determined the AV vector for the frame as that vector having a maximum correlation value wherein the correlation values are determined from a subset of the audio and video features.

Amendment  
Serial No. 10/076,194

Docket No. US020037

A claimed invention is prima facie obvious when three basic criteria are met. First, there must be some suggestion or motivation, either in the reference themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine the teachings therein. Second, there must be a reasonable expectation of success. And, third, the prior art reference or combined references must teach or suggest all the claim limitations.

In this case, Basu, in the alternate embodiment shown in Figure 3, teaches determining a single AV vector based on the combined audio and visual features. However, Basu fails to teach or suggest the determination of the vector as described in the claims. Navenda similarly is silent with regard to the manner of determining the AV vector and is cited for teaching that audio component has certain features.

Thus, even if the teachings of Basu and Navenda were combined as suggested by the Office Action, the combined device would not render obvious the present invention as the combined device fails to recite all the elements claims.

For at least this reason, applicant submits that the reason for the rejections of claim 1 has been overcome and the rejection can no longer be sustained. Applicant respectfully requests withdrawal of the rejection and allowance of the claims.

With regard to the remaining independent claims, these claims recite subject matter similar to that recited in claim 1 and were rejected citing the same references used in rejecting claim 1. Thus, applicant's remarks made in response to the rejection of claim 1 are also applicable in response to the rejection of the remaining independent claims. Applicant submits that in view of the remarks made with regard to the rejection of claim 1, which are reasserted, as if in full, in response to the rejection of the remaining independent claims, the reason for the rejection of these claims has been overcome and the rejection can no longer be sustained. Applicant respectfully requests withdrawal of the rejection and allowance of the claims.

With regard to remaining dependent claims, these claims ultimately depend from the independent claims, which has been shown not to be obvious, and, hence, allowable, over the cited references. Accordingly, the aforementioned dependent claims are also allowable by virtue of their dependence from an allowable base claim.

Amendment  
Serial No. 10/076,194

Docket No. US020037

Claims 6-7 stand rejected under 35 USC 103(a) as being unpatentable over Basu in view of Nevenka and further in view of Bradford (USPPA 2002/0103799). Claims 9-10, 12-14 and 18-20 stand rejected under 35 USC 103(a) as being unpatentable over Basu in view of Nevenka and further in view of Wang (Multimedia Content Analysis).

Applicant respectfully disagrees with and explicitly traverses the reason for rejecting the aforementioned claims. These claims depend from the independent claims, which has been shown to contain subject matter not disclosed by the combination of Basu and Nevenda. The additionally cited references fail to provide any teaching or suggestion to correct the deficiency noted in the combination of the primary references. Hence, even if there were some motivation to combine the teachings of all of the cited references, the device so created fails to teach all the features recited in the independent claims, and consequently, the aforementioned dependent claims.

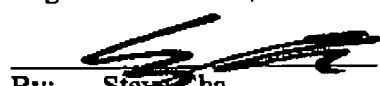
Accordingly, the invention recited in the aforementioned claims is not rendered obvious by the teachings of the cited references. For at least this reason applicant submits that the reason for the rejection has been overcome and respectfully requests that the rejection be withdrawn.

For all the foregoing reasons, it is respectfully submitted that all the present claims are patentable in view of the cited references. A Notice of Allowance is respectfully requested.

Respectfully submitted,

Dan Piotrowski  
Registration No. 42,079

Date: August 24, 2006

  
By: Steven Cha  
Attorney for Applicant  
Registration No. 44,069

**Mail all correspondence to:**  
Dan Piotrowski, Registration No. 42,079  
US PHILIPS CORPORATION  
P.O. Box 3001  
Briarcliff Manor, NY 10510-8001  
Phone: (914) 333-9624  
Fax: (914) 332-0615